

AD-A085 207

SOUTHERN METHODIST UNIV DALLAS TEX DEPT OF STATISTICS

F/6 12/1

L-STATISTICS WITH SMOOTH WEIGHT FUNCTIONS JACKKNIFE WELL.(U)

MAY 80 W C PARR, W R SCHUCANY

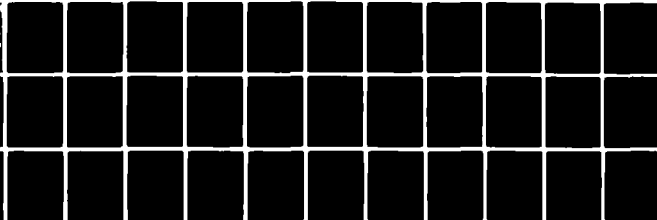
N00014-75-C-0439

UNCLASSIFIED

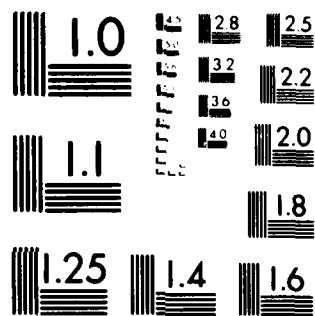
TR-138

NL

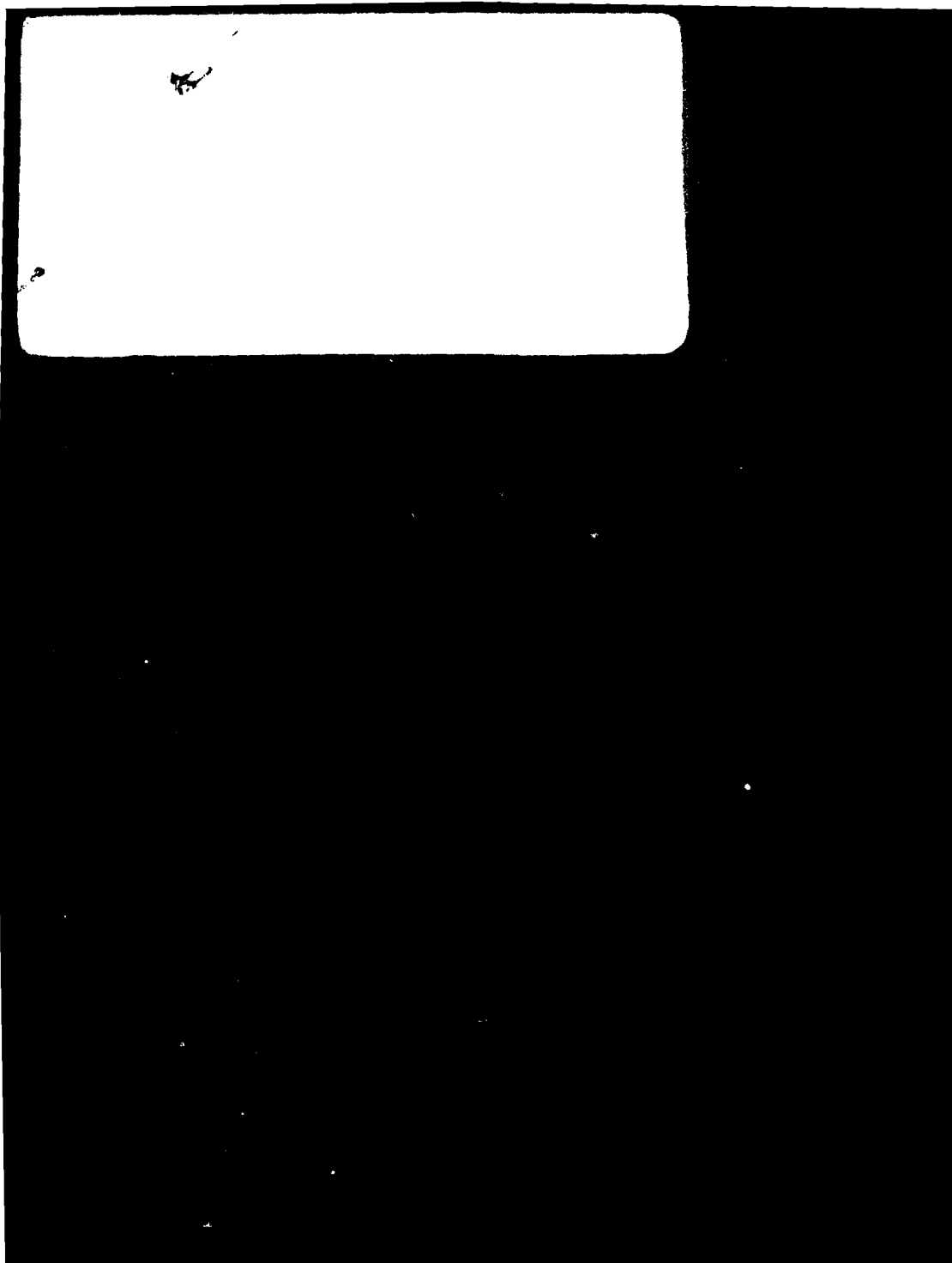
1-1
A 40007



END
DATE
FILMED
6 80
DTIC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A



L-STATISTICS WITH SMOOTH WEIGHT

FUNCTIONS JACKKNIFE WELL

by

William C. Parr
William R. Schucany

Technical Report No. 138
Department of Statistics ONR Contract
May 13, 1980

Research Sponsored by the Office of Naval Research
Contract N00014-75-C-0439 /
Project NR 042-280

Reproduction in whole or in part is permitted
for any purpose of the United States Government

This document has been approved for public release
and sale; its distribution is unlimited

JUN 9 1980

A

DIVISION OF MATHEMATICAL SCIENCES
Department of Statistics
Southern Methodist University
Dallas, Texas 75275

Author	
Editor	
Reviewer	
Project Manager	
Other	
Date	
Initials	
Signature	
Title	
Institution	
Address	
City	
State	
Zip	

A

L-STATISTICS WITH SMOOTH WEIGHT
FUNCTIONS JACKKNIFE WELL

by

William C. Parr and William R. Schucany

ABSTRACT

↙
The behavior of linear combinations of order statistics (L-statistics) under jackknifing is discussed. The asymptotic behavior of a jackknifed L-statistic is produced, and the pseudo-value based variance estimate is shown to be consistent under moderate smoothness and a trimming condition of the weight function. The results extend easily to functions of L-statistics and thus answer a question posed in Miller (1974). Monte Carlo results support small-sample applicability of the large-sample results for the construction of approximate confidence intervals.

Key Words: Confidence intervals; Jackknife; Linear combinations of order statistics; Variance estimation.

William C. Parr is Assistant Professor, Institute of Statistics, Texas A&M University, College Station, Texas 77843. William R. Schucany is Professor of Statistics, Southern Methodist University, Dallas, Texas 75275. Research of the second author was supported in part by ONR Contract N00014-75-C-0439. The authors would like to thank Emanuel Parzen for helpful conversations.

1

1. INTRODUCTION, NOTATION, AND DEFINITIONS

Let X_1, X_2, \dots, X_n be a random sample of size n from a distribution with distribution function F , and let $X_{1,n}, X_{2,n}, \dots, X_{n,n}$ denote the associated order statistics. For a fixed weight function $J(u)$ defined for $0 < u < 1$, we define the L-statistic

$$S_n = \frac{1}{n} \sum_{i=1}^n J\left(\frac{i}{n+1}\right) X_{i,n}. \quad (1.1)$$

Other definitions, typically asymptotically equivalent to that

above, include $T_n = \frac{1}{n} \sum_{i=1}^n J\left(\frac{i}{n}\right) X_{i,n}$ and $U_n = \frac{1}{n} \sum_{i=1}^n c_{i,n} X_{i,n}$

where $c_{i,n} = \int_{(i-1)/n}^{i/n} J(u) du$. S_n is chosen for study in this paper

as being typical of actual L-statistics used in practice (use of U_n , in which the integration "smooths" J , might result in fewer conditions on the weight function).

L-statistics of the form of S_n are often used in estimation problems since they are typically computationally simple and (at least for location and scale problems) asymptotically efficient given the proper choice of the weight function J . Thus, they are often good choices a) as estimates for their own sake, b) as good starting values for iterative estimation

procedures, and c) as quick and consistent estimators of nuisance parameters (such as unknown scale in regression problems) to minimize the number of parameters being simultaneously estimated via an iterative procedure. Herein we consider primarily a), where L-statistics are used to make parametric inferences on their own. However, S_n is often biased as an estimator of $S_0 = \int_0^1 Q(u)J(u)du$ where $Q(u) = \inf\{x:F(x) \geq u\}$, as is $g(S_n)$ biased for $g(S_0)$ for many choices of g . Also, there seems to be a dearth of procedures for consistent nonparametric estimation of the variance of $\sqrt{n}(S_n - S_0)$. (But see Sen (1979) in this regard.) For both of these problems, reduction of bias and consistent nonparametric variance estimation, the jackknife is a natural choice as a possibly non-optimal rough-and-ready tool.

The ordinary jackknife of S_n may be written as

$$\begin{aligned}
 \tilde{S}_n &= \frac{1}{n} \sum_{i=1}^n S_{i,n} \\
 &= \sum_{i=1}^n J\left(\frac{1}{n+1}\right) X_{i,n} \\
 &\quad - \frac{1}{n} \sum_{i=1}^n \left[(i-1)J\left(\frac{i-1}{n}\right) + (n-i)J\left(\frac{i}{n}\right) \right] X_{i,n} \\
 &= n S_n - \frac{n-1}{n} \sum_{i=1}^n S_{n-1}^{(i)} \quad ,
 \end{aligned} \tag{1.3}$$

where the i th pseudo-value is

$$S_{i,n} = \sum_{j=1}^n J\left(\frac{j}{n+1}\right) X_{j,n} - \sum_{j=1}^{i-1} J\left(\frac{j}{n}\right) X_{j,n} - \sum_{j=i+1}^n J\left(\frac{j-1}{n}\right) X_{j,n}, \quad (1.4)$$

$S_{n-1}^{(i)}$ is the same L-statistic computed after deleting $X_{i,n}$ from the sample, and $\sum_{i=a}^b h(i)$ is understood to be defined to be zero if $a > b$. (See Miller (1974) for basic definitions of the jackknife and pseudo-values in general.) Note that the definition of J at 0 and 1 is completely arbitrary since the associated terms will cancel out in (1.3).

We further define the sample variance of the pseudo-values as

$$s_p^2 = \frac{1}{n-1} \sum_{i=1}^n (S_{i,n} - \bar{S}_n)^2. \quad (1.5)$$

2. RESULTS ON JACKKNIFING L-STATISTICS

Theorem 1: Let X_1, X_2, \dots, X_n be a random sample of size n from a distribution F with $E_F |X|^p < \infty$ for some $\frac{1}{2} \leq p \leq 1$.

Further, let S_n and \bar{S}_n be as defined by (1.1) and (1.3). If

- a) $p = 1$ and J' satisfies a Hölder condition with $\alpha > \frac{1}{2}$, or
- b) $\frac{1}{2} \leq p < 1$ and J' satisfies a Hölder condition with

$\alpha = \frac{1}{p} - \frac{1}{2}$, then $\sqrt{n}(\tilde{S}_n - S_n) \rightarrow 0$ with probability 1 as $n \rightarrow \infty$.

Proof: From elementary algebra, $\tilde{S}_n - S_n = \sum_{i=1}^n D_J(i,n) X_{i,n}$,
where

$$D_J(i,n) = \frac{n-1}{n} J\left(\frac{i}{n+1}\right) - \frac{i-1}{n} J\left(\frac{i-1}{n}\right) - \frac{n-i}{n} J\left(\frac{i}{n}\right),$$

$$1 \leq i \leq n.$$

Since J' obeys a Hölder condition with exponent α we have

(with $\frac{1}{n+1} \leq \xi_{in1} \leq \frac{i}{n}$)

$$\begin{aligned} J\left(\frac{i}{n}\right) &= J\left(\frac{i}{n+1}\right) + J'\left(\frac{i}{n+1}\right)\left(\frac{i}{n} - \frac{i}{n+1}\right) \\ &\quad + (J'(\xi_{in1}) - J'\left(\frac{i}{n+1}\right))\left(\frac{i}{n} - \frac{i}{n+1}\right) \\ &= J\left(\frac{i}{n+1}\right) + J'\left(\frac{i}{n+1}\right) \frac{i}{n(n+1)} + B_{in1} \left|\frac{i}{n(n+1)}\right|^{1+\alpha} \end{aligned}$$

and (with $\frac{i-1}{n} < \xi_{in2} < \frac{i}{n+1}$)

$$\begin{aligned}
J\left(\frac{i-1}{n}\right) &= J\left(\frac{i}{n+1}\right) + J'\left(\frac{i}{n+1}\right)\left(\frac{i-1}{n} - \frac{i}{n+1}\right) \\
&\quad + (J'(\xi_{in2}) - J'\left(\frac{i}{n+1}\right))\left(\frac{i-1}{n} - \frac{i}{n+1}\right) \\
&= J\left(\frac{i}{n+1}\right) + J'\left(\frac{i}{n+1}\right) \frac{i-n-1}{n(n+1)} \\
&\quad + B_{in2} \left| \frac{i-n-1}{n(n+1)} \right|^{1+\alpha},
\end{aligned}$$

where the B_{in1} and B_{in2} are all uniformly bounded in absolute value by some finite positive constant B . Hence,

$$\begin{aligned}
D_J(i,n) &= -J'\left(\frac{i}{n+1}\right) \frac{n-2i+1}{n^2(n+1)} - \frac{n-i}{n} B_{in1} \left(\frac{i}{n(n+1)}\right)^{1+\alpha} \\
&\quad - \frac{i-1}{n} B_{in2} \left| \frac{i-n-1}{n(n+1)} \right|^{1+\alpha}.
\end{aligned}$$

Thus,

$$\begin{aligned}
\sqrt{n} |\bar{S}_n - S_n| &\leq \sqrt{n} \sum_{i=1}^n |D_J(i,n)| |X_{i,n}| \\
&\leq n^{1+\alpha} \sup_{1 \leq i \leq n} |D_J(i,n)| \sum_{i=1}^n \frac{|X_i|}{n^{(1+2\alpha)/2}} \\
&= Z_n.
\end{aligned}$$

If $0 < \frac{2}{1+2\alpha} < 1$ and $E_F |X|^{\frac{2}{1+2\alpha}} < \infty$, with J' satisfying a Hölder condition with exponent α , $Z_n \rightarrow 0$ with probability one, by Marcinkiewicz' theorem (Loeve (1977), p.254). If $E_F |X| < \infty$ and J' satisfies a Hölder condition with exponent $\alpha > \frac{1}{2}$, $Z_n \rightarrow 0$ with probability one, using the strong law of large numbers.

Several comments are in order:

1) This result is the univariate analogue for L-statistics of result J.1 of Reeds (1978) for M-estimators. The extension to the multivariate case is immediate and hence omitted.

2) The moment condition on F is seen by an inspection of the proof to be superfluous if S "trims", i.e., if $J(u) = 0$ for $u \in (0, \epsilon) \cup (1 - \epsilon, 1)$ for some $0 < \epsilon < \frac{1}{2}$.

3) If further $\sqrt{n}(S_n - S_0) \xrightarrow{d} N(0, \sigma^2)$, where

$$\sigma^2 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} J(F(x))J(F(y)) [F(\min(x,y)) - F(x)F(y)] dx dy > 0, \quad (2.1)$$

then $\sqrt{n}(\tilde{S}_n - S_0) \xrightarrow{d} N(0, \sigma^2)$ likewise. This is true if the conditions of Theorem 1 hold with $p = 1$ and J' is of bounded

variation (using the result of D. S. Moore (1968), with $\sigma^2 < \infty$).

4) Since J'' bounded on $(0,1)$ implies that J' satisfies a Hölder condition with $\alpha = 1$, the condition on J and F in the theorem may be replaced by the stronger but intuitively clearer condition that J'' is bounded and $E_F |X|^{2/3} < \infty$.

5) A similar theorem was stated under much stronger conditions by Thorburn (1976).

6) Theorem 1 yields a law of the iterated logarithm for the jackknife of a linear function of order statistics from the corresponding law for the original statistics. Using Theorem 4 (Example 1) of Wellner (1977b), we obtain that if $E_F |X|^{2+\epsilon} < \infty$ for some $\epsilon > 0$ (and $\alpha > \frac{1}{2}$)

$$\limsup_n \frac{\sqrt{n} \left| \tilde{S}_n - \int_0^1 Q(u) J_n(u) du \right|}{\sqrt{2 \sigma^2 \log \log n}} = 1$$

with probability one, where $J_n(u) = J\left(\frac{i}{n+1}\right)$ for $\frac{i-1}{n} < u \leq \frac{i}{n}$, $1 \leq i \leq n$, and $J_n(0) = J\left(\frac{1}{n+1}\right)$.

7) Similarly, a Berry-Esseen rate for \tilde{S}_n follows directly from that of Helmers (1977). If $E_F |X|^3 < \infty$, $\int_0^1 |J'(u)| dQ(u) < \infty$ and $\sigma^2 > 0$, we quickly obtain that

$$\sup_x |F_n^*(x) - \phi(x)| = O(n^{-1/2}),$$

where $F_n^*(x)$ is the cumulative distribution of $(\tilde{S}_n - E[S_n]) / (\text{Var} S_n)^{1/2}$ and $\phi(\cdot)$ is the standard normal cumulative.

The following theorem gives conditions under which the jackknife provides a consistent estimator of the asymptotic variance of $\sqrt{n}(S_n - S_0)$ and $\sqrt{n}(\tilde{S}_n - S_0)$, and hence makes possible the construction of asymptotically pivotal quantities.

Theorem 2: Let X_1, \dots, X_n be a random sample of size n from a distribution F , and let s_p^2 be defined by (1.5). If there exist positive numbers ϵ and δ such that $0 < \delta < \epsilon < \frac{1}{2}$, (let $C(\epsilon) = [0, \epsilon] \cup (1 - \epsilon, 1]$; $J(u) = 0$ for $u \in C(\epsilon)$, $fQ(u) \geq B > 0$ for $u \in C(\epsilon - \delta)$, and J' satisfies a Hölder condition with exponent $\alpha > 0$, then $s_p^2 \xrightarrow{P} \sigma^2$, with σ^2 given by (2.1).

Proof: Note that $\sigma^2 = \text{Var}(H(U))$ where $U \sim u(0,1)$ and $H(u) = Q(u)J(u) - \int_0^1 Q(t)J'(t)[t - I(u \leq t)]dt$, translating

the results of Boos (1979, eq.3.3) into the quantile domain and then integrating by parts. Now, from (1.5), it will suffice for the desired result if we show that

$$\sup_{1 \leq i \leq n} |S_{i,n} - H\left(\frac{i}{n+1}\right)| \xrightarrow{P} 0, \quad \text{as } n \rightarrow \infty. \quad (2.2)$$

If (2.2) holds, then $s_p^2 = \frac{1}{n-1} \sum_{i=1}^n \left(H\left(\frac{i}{n+1}\right) - \frac{1}{n} \sum_{j=1}^n H\left(\frac{j}{n+1}\right) \right)^2 + o_p(1)$.

Furthermore, $\frac{1}{n-1} \sum_{i=1}^n H^2\left(\frac{i}{n+1}\right)$ and $\frac{1}{n} \sum_{i=1}^n H\left(\frac{i}{n+1}\right)$ will

converge to $\int_0^1 H^2(u) du$ and $\int_0^1 H(u) du$ respectively, giving

$s_p^2 \xrightarrow{P} \sigma^2$. By definition,

$$\begin{aligned} S_{i,n} &= \sum_{j=1}^n J\left(\frac{j}{n+1}\right) X_{j,n} - \sum_{j=1}^{i-1} J\left(\frac{j}{n}\right) X_{j,n} - \sum_{j=i+1}^n J\left(\frac{j-1}{n}\right) X_{j,n} \\ &= J\left(\frac{i}{n+1}\right) X_{i,n} - \sum_{j=1}^{i-1} \left\{ J\left(\frac{j}{n}\right) - J\left(\frac{j}{n+1}\right) \right\} X_{j,n} \\ &\quad - \sum_{j=i+1}^n \left\{ J\left(\frac{j-1}{n}\right) - J\left(\frac{j}{n+1}\right) \right\} X_{j,n} \\ &= J\left(\frac{i}{n+1}\right) X_{i,n} - \frac{1}{n} \sum_{j=1}^{i-1} J'\left(\frac{j}{n+1}\right) \frac{j}{n+1} X_{j,n} - \frac{1}{n} \sum_{j=i+1}^n J'\left(\frac{j}{n+1}\right) \\ &\quad \left(\frac{j}{n+1} - 1 \right) X_{j,n} + R_{n1i} \end{aligned}$$

$$= J\left(\frac{1}{n+1}\right)X_{1,n} - \frac{1}{n} \sum_{j=1}^n \left(J'\left(\frac{1}{n+1}\right) \frac{1}{n+1} - I\left(\frac{1}{n+1} \leq \frac{1}{n+1}\right) \right) X_{j,n} \\ + R_{n21},$$

using Taylor expansions and the Hölder condition on J' . Also using trimming condition on J , the remainder term is such that

$$\sup_{1 \leq i \leq n} |R_{n21}| = o_p(1) \text{ as } n \rightarrow \infty.$$

From the continuity and boundedness of J and the fact

$$\text{that } fQ(u) \geq B > 0 \text{ wherever } J(u) \neq 0, \quad J\left(\frac{1}{n+1}\right)X_{1,n} = \\ J\left(\frac{1}{n+1}\right)Q\left(\frac{1}{n+1}\right) + R_{n31}, \text{ with } \sup_{1 \leq i \leq n} |R_{n31}| = o_p(1). \text{ Further}$$

$$\frac{1}{n} \sum_{j=1}^n J'\left(\frac{1}{n+1}\right) \frac{1}{n+1} X_{j,n} \rightarrow \int_0^1 Q(t)J'(t)F(t)dt \text{ with probability}$$

one from Corollary 2 to Theorem 1 of Wellner (1977).

Then, for fixed K let $u_i = \frac{1}{K+1}$, $i = 0, 1, \dots, K+1$.

It follows easily that

$$L_1 = \frac{1}{n} \sum_{j=[nu_1]}^n J'\left(\frac{1}{n+1}\right)X_{j,n} \xrightarrow{P} \int_{u_1}^1 J'(t)Q(t)dt \\ = \int_0^1 J'(t)I(u_1 \leq t)Q(t)dt.$$

Since K is finite, the convergence in probability is uniform in i , $i = 0, 1, \dots, K+1$. Furthermore, letting $\inf_{i=1, \dots, K} |u_i - u|$ be achieved at index value i^* ,

$$\begin{aligned} & \sup_{0 \leq u \leq 1} \left| L_{i^*} - \frac{1}{n} \sum_{j=[nu]}^n J' \left(\frac{j}{n+1} \right) X_{j,n} \right| \\ & \leq \sup_{0 \leq u \leq 1} |J'(u)| \frac{|F_n^{-1}(\epsilon)| + |F_n^{-1}(1 - \epsilon)|}{K+1}, \end{aligned}$$

which may be made less than any specified positive number in probability through a sufficiently large choice of K . Hence, (2.2) holds and $s_p^2 \xrightarrow{P} \sigma^2$, by the uniform continuity of $\int_u^1 J'(t)Q(t)dt$.

Some pertinent comments follow:

1) This result provides a method for consistent variance estimation for L -statistics, being the univariate analogue for L -statistics of result J.2 of Reeds (1978) for M -estimators.

2) Finiteness of σ^2 is clearly implied by the trimming and boundedness conditions on J .

3) This result makes possible the construction of nonparametric approximate confidence intervals for

$$S_0 = \int_0^1 Q(u)J(u)du, \text{ using as pivotal quantities}$$

$$\frac{\sqrt{n} (\bar{S}_n - S_0)}{s_p} \quad \text{or} \quad \frac{\sqrt{n} (S_n - S_0)}{s_p}.$$

This is, to the knowledge of the authors, the only nonparametric method of consistent variance estimation for L-statistics (i.e., in the absence of a specified parametric form for the unknown density) other than that of Sen (1979) discussed below. For a specific parametric family, a consistent estimator would of course typically be provided by

$$\hat{\sigma}^2 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} J(\hat{F}_{\hat{\theta}}(x)) J(\hat{F}_{\hat{\theta}}(y)) [\hat{F}_{\hat{\theta}}(\min(x,y)) - \hat{F}_{\hat{\theta}}(x) \hat{F}_{\hat{\theta}}(y)] dx dy, \quad (2.3)$$

where F_{θ} , $\theta \in \Omega$ is the parametric family of densities and $\hat{\theta}$ is a weakly consistent estimator of θ , possibly multivariate.

4) It would be interesting to compare the properties of this estimator with that proposed by Sen (1979), which is essentially

$$\tilde{\sigma}^2 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} J(F_n(x)) J(F_n(y)) [F_n(x,y) - F_n(x) F_n(y)] dx dy, \quad (2.4)$$

obtained by substituting the empirical distribution function F_n for F in (2.1).

5) Relaxation of the trimming condition on J , which would require either moment conditions on F or joint conditions governing $J(u)$ and $Q(u)$ as u approaches 0 and 1 is questionable from

the standpoint of robust inference and hence Theorem 2 is stated as being of interest in its own right. Theorem 3, however, obtains strong convergence of s_p^2 , while dropping the trimming condition, at the cost of moment assumptions.

Theorem 3: Let X_1, \dots, X_n be a random sample of size n from a distribution F with $E|X|^{2+\epsilon} < \infty$ for some $\epsilon > 0$. If J' is continuous on $[0,1]$, then $s_p^2 \rightarrow \sigma^2$ with probability one.

Proof: We proceed as in the proof of Theorem 2 to write

$$S_{i,n} = J\left(\frac{j}{n+1}\right)X_{j,n} - \frac{1}{n} \sum_{j=1}^n J'\left(\frac{j}{n+1}\right) \left(\frac{j}{n+1} - I(j \leq 1)\right) X_{j,n} \\ - \frac{1}{n} \sum_{j=1}^n (J'(\eta_{nji}) - J'\left(\frac{j}{n+1}\right)) \left(\frac{j}{n+1} - I(j \leq 1)\right) X_{j,n} ,$$

where $\frac{j-1}{n} \leq \eta_{nji} \leq \frac{j}{n}$ and in particular $\eta_{nii} = \frac{i}{n+1}$.

The proof is then concluded by showing i) $\frac{1}{n} \sum_{i=1}^n S_{i,n} \rightarrow \int_0^1 Q(u)J(u)du$,

and ii) $\frac{1}{n} \sum_{i=1}^n S_{i,n}^2 \rightarrow \sigma^2 + \left(\int_0^1 Q(u)J(u)du\right)^2$, both convergences

being with probability one. Proceeding with the first part,

$$\begin{aligned} \frac{1}{n} \sum_{i=1}^n S_{i,n} &= \frac{1}{n} \sum_{i=1}^n J\left(\frac{i}{n+1}\right) X_{i,n} - \frac{1}{n} \sum_{j=1}^n J'\left(\frac{j}{n+1}\right) \left(\frac{j}{n+1} - \frac{j}{n}\right) X_{j,n} \\ &\quad - \frac{1}{n} \sum_{i=1}^n \frac{1}{n} \sum_{j=1}^n (J'(\eta_{nji}) - J'\left(\frac{j}{n+1}\right)) \left(\frac{j}{n+1} - I(j \leq i)\right) X_{j,n}. \end{aligned}$$

The first term converges to $\int_0^1 Q(u)J(u)du$ with probability one by Theorem 4, example 1 of Wellner (1977a).

The second term is $-\frac{1}{n} \left\{ \frac{1}{n} \sum_{j=1}^n J'\left(\frac{j}{n+1}\right) \frac{j}{n+1} X_{j,n} \right\}$ which

converges to 0 with probability one by the same result (since

$$\frac{1}{n} \sum_{i=1}^n J'\left(\frac{j}{n+1}\right) \frac{j}{n+1} X_{j,n} \text{ converges to}$$

$\int_0^1 Q(u)uJ'(u)du$ with probability one). Lastly,

$$\begin{aligned} &\left| \frac{1}{n} \sum_{i=1}^n \frac{1}{n} \sum_{j=1}^n (J'(\eta_{nji}) - J'\left(\frac{j}{n+1}\right)) \left(\frac{j}{n+1} - I(j \leq i)\right) X_{j,n} \right| \\ &\leq \sup_{i,j} |J'(\eta_{nji}) - J'\left(\frac{j}{n+1}\right)| \frac{\sum_{j=1}^n |X_{j,n}|}{n} \end{aligned}$$

which converges to zero with probability one using the uniform continuity of J' and the first moment assumption on F . For the second part,

$$\begin{aligned}
\frac{1}{n} \sum_{i=1}^n S_{i,n}^2 &= \frac{1}{n} \sum_{i=1}^n \left[J\left(\frac{1}{n+1}\right) X_{i,n} - \frac{1}{n} \sum_{j=1}^n J'\left(\frac{1}{n+1}\right) \left\{ \frac{1}{n+1} - I(j \leq i) \right\} X_{j,n} \right. \\
&\quad \left. - \frac{1}{n} \sum_{j=1}^n (J'(\eta_{nj1}) - J'\left(\frac{1}{n+1}\right)) \left\{ \frac{1}{n+1} - I(j \leq i) \right\} X_{j,n} \right]^2 \\
&= \frac{1}{n} \sum_{i=1}^n \left[J\left(\frac{1}{n+1}\right) X_{i,n} - \frac{1}{n} \sum_{j=1}^n J'\left(\frac{1}{n+1}\right) \left\{ \frac{1}{n+1} - I(j \leq i) \right\} X_{j,n} \right]^2 \\
&\quad - 2 \frac{1}{n} \sum_{i=1}^n \left[J\left(\frac{1}{n+1}\right) X_{i,n} - \frac{1}{n} \sum_{j=1}^n J'\left(\frac{1}{n+1}\right) \left\{ \frac{1}{n+1} - I(j \leq i) \right\} X_{j,n} \right] \cdot \\
&\quad \frac{1}{n} \sum_{j=1}^n (J'(\eta_{nj1}) - J'\left(\frac{1}{n+1}\right)) \left\{ \frac{1}{n+1} - I(j \leq i) \right\} X_{j,n} \\
&\quad + \frac{1}{n} \sum_{i=1}^n \left\{ \frac{1}{n} \sum_{j=1}^n (J'(\eta_{nj1}) - J'\left(\frac{1}{n+1}\right)) \left\{ \frac{1}{n+1} - I(j \leq i) \right\} X_{j,n} \right\}^2 \\
&= A_{1n} + A_{2n} + A_{3n} .
\end{aligned}$$

The third term, A_{3n} , is less than or equal to (in absolute value)

$$\left(\frac{1}{n} \sup_{i,j} |J'(\eta_{nj1}) - J'\left(\frac{1}{n+1}\right)| \sum_{i=1}^n |X_i| \right)^2 ,$$

which converges with probability one to zero via the strong law of large numbers and the continuity condition on J' . The

second term, A_{2n} , is disposed of in similar fashion. The first term, A_{1n} , requires a more extended analysis.

$$\begin{aligned}
 A_{1n} &= \frac{1}{n} \sum_{i=1}^n J^2 \left(\frac{i}{n+1} \right) X_{i,n}^2 + \frac{1}{n} \sum_{i=1}^n \left[\frac{1}{n} \sum_{j=1}^n J' \left(\frac{i}{n+1} \right) \left\{ \frac{i}{n+1} - I(j \leq i) \right\} X_{j,n} \right]^2 \\
 &\quad - \frac{2}{n} \sum_{i=1}^n \left\{ J \left(\frac{i}{n+1} \right) X_{i,n} \cdot \frac{1}{n} \sum_{j=1}^n J' \left(\frac{i}{n+1} \right) \left\{ \frac{i}{n+1} - I(j \leq i) \right\} X_{j,n} \right\} \\
 &= B_{1n} + B_{2n} + B_{3n} .
 \end{aligned}$$

$B_{1n} \rightarrow \int_0^1 Q^2(u) J^2(u) du$ with probability one by Wellner's Theorem 4. By a similar argument and rearrangement of terms,

$$\begin{aligned}
 B_{2n} &\rightarrow \left(\int_0^1 Q(u) J'(u) u du \right)^2 - 2 \int_0^1 u J'(u) Q(u) du \cdot \int_0^1 J'(u) (1-u) Q(u) du \\
 &\quad + \int_0^1 \left(\int_0^u Q(v) J'(v) dv \right)^2 du
 \end{aligned}$$

and

$$\begin{aligned}
 B_{3n} &\rightarrow -2 \int_0^1 Q(u) J(u) du \cdot \int_0^1 Q(u) J'(u) u du \\
 &\quad - 2 \int_0^1 Q(u) J(u) \left(\int_0^u Q(v) J'(v) dv \right) du ,
 \end{aligned}$$

both convergences holding with probability one. The result then follows by integrating the expression for σ^2 given by (2.1) by parts and observing that the quantity to which s_p^2 converges with probability one is indeed σ^2 . (Note that the appeals to Wellner's result are actually to a slight modification of it allowing random J_n which satisfy the boundedness and convergence conditions with probability one.)

Example 1: The sample mean, for which $J(u) \equiv 1$, clearly satisfies the conditions of Theorem 1 (but not the trimming conditions of Theorem 2) if $E_F|X|^{2/3} < \infty$. In fact, $\bar{S}_n \equiv S_n$

in this case, and $s_p^2 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n-1} \xrightarrow{P} \sigma_F^2$ if $E_F|X|^2 < \infty$.

Theorem 3 requires $E_F|X|^{2+\epsilon} < \infty$ for some $\epsilon > 0$. Thus, the usual strong law for the sample variance "just fails" to be a corollary of Theorem 3.

Example 2: While the ordinary trimmed means do not satisfy the score conditions of Theorems 1 and 2, a smoothed version causing J to return to zero in such a way that it is differentiable with J' obeying a Hölder condition with $\alpha > \frac{1}{2}$ would satisfy those conditions. We assume that the modified score is also zero on $C(\epsilon)$.

Example 3: Gini's mean difference ($J(u) = u - \frac{1}{2}$) and the optimal score for location estimation for a logistic population ($J(u) = 6u(1 - u)$) clearly satisfy the conditions of Theorem 1 if $E_F(X) < \infty$ and those of Theorem 3 if $E|X|^{2+\epsilon} < \infty$, but violate the trimming conditions of Theorem 2.

In most instances, Theorems 2 or 3 will be of primary interest, providing methods for the construction of approximate confidence intervals based upon L-statistics (and one-to-one functions thereof). The bias of an L-statistic will often be small (but see Section 3 in this regard), and hence \bar{S}_n will be of limited practical use for the purposes of estimation with reduced bias. However, the end goal of an analysis may be to estimate or construct approximate tests or confidence intervals for $g(S_0)$. For estimation, $g(S_n)$ may suffer from severe bias if g is highly non-linear near S_0 (recall that

$E_F[g(S_n)] \approx g(S_0) + \frac{g''(S_0)}{2} \text{Var}_F(S_n)$). Hence jackknifing $g(S_n)$ would be of interest in such cases for reduction of bias.

If g is non-monotone, confidence intervals for $g(S_n)$ obtained by finding a confidence interval for S_0 using $\sqrt{n}(S_n - S_0)/s_p$ as an approximately standard normal pivotal quantity and taking the image of such an interval under g may well result in longer intervals than would be obtained by pivoting about $g(S_n)$.

Thus, it seems to be of independent interest to study the behavior of $g(S_n)$ under jackknifing. The following theorem parallels Theorem 1, establishing that $g(S_n)$ and its jackknife have the same limiting distribution.

Theorem 4: Let X_1, X_2, \dots, X_n be a random sample of size n from a distribution F with $E[|X|^{4/3}] < \infty$. Let J' obey a Hölder condition with $\alpha > \frac{1}{2}$. If g is a function with a bounded second derivative in a neighborhood of S_0 , and $\tilde{g}(S_n) =$

$ng(S_n) - \frac{n-1}{n} \sum_{i=1}^n g(S_{n-1}^{(i)})$, the jackknife of $g(S_n)$, then

$$\sqrt{n}(g(S_n) - \tilde{g}(S_n)) \rightarrow 0 \text{ with probability one.}$$

Proof:

$$\begin{aligned} \sqrt{n}(g(S_n) - \tilde{g}(S_n)) &= \sqrt{n} \frac{(n-1)}{n} \sum_{i=1}^n [g'(S_n)(S_{n-1}^{(i)} - S_n) \\ &\quad + \frac{1}{2} g''(\xi_{in})(S_{n-1}^{(i)} - S_n)^2] \end{aligned}$$

$$\begin{aligned}
 &= \frac{n-1}{n} g'(S_n) \sqrt{n} (\bar{S}_n - S_n) \\
 &\quad + \frac{n-1}{2n} \sqrt{n} \sum_{i=1}^n (g''(\xi_{in})) (S_{n-1}^{(i)} - S_n)^2,
 \end{aligned} \tag{2.5}$$

where ξ_{in} is between S_n and $S_{n-1}^{(i)}$. To justify this expansion, we need to prove

$$\sup_{1 \leq i \leq n} |S_{n-1}^{(i)} - S_n| \rightarrow 0 \text{ with probability one.}$$

Now,

$$\begin{aligned}
 S_{n-1}^{(i)} - S_n &= \sum_{j=1}^{i-1} \left[\frac{1}{n-1} J\left(\frac{j}{n}\right) - \frac{1}{n} J\left(\frac{j}{n+1}\right) \right] X_{j,n} \\
 &\quad + \sum_{j=i+1}^n \left[\frac{1}{n-1} J\left(\frac{j-1}{n}\right) - \frac{1}{n} J\left(\frac{j}{n+1}\right) \right] X_{j,n} \\
 &\quad - \frac{1}{n} J\left(\frac{i}{n+1}\right) X_{i,n} \\
 &= \sum_{\substack{j=1 \\ j \neq i}}^n \left[\frac{1}{n-1} \left\{ J\left(\frac{j}{n+1}\right) + J'\left(\frac{j}{n+1}\right) \frac{j}{n(n+1)} \right. \right. \\
 &\quad \left. \left. + (J'(n_{nj}) - J'\left(\frac{j}{n+1}\right)) \frac{j}{n(n+1)} \right\} \right. \\
 &\quad \left. - \frac{1}{n} J\left(\frac{j}{n+1}\right) \right] X_{j,n} - \frac{1}{n} J\left(\frac{i}{n+1}\right) X_{i,n}
 \end{aligned}$$

$$\begin{aligned}
&= \sum_{\substack{j=1 \\ j \neq i}}^n \left\{ \frac{1}{n(n-1)} J\left(\frac{j}{n+1}\right) + \frac{1}{n(n+1)(n-1)} J'\left(\frac{j}{n+1}\right) \right. \\
&\quad \left. + \frac{1}{n(n+1)(n-1)} (J'(\eta_{nj}) - J'\left(\frac{j}{n+1}\right)) \right\} X_{j,n} \\
&\quad - \frac{1}{n} J\left(\frac{i}{n+1}\right) X_{i,n} .
\end{aligned}$$

So

$$\sup_{1 \leq i \leq n} |S_{n-1}^{(i)} - S_n| \leq \sum_{j=1}^n \frac{M}{n^2} |X_{j,n}| + \frac{1}{n} \sup_{0 \leq u \leq 1} |J(u)| (|X_{n,n}| + |X_{1,n}|)$$

with M determined by the bounds on J and J' . Finally, this expression converges to zero with probability one by the strong law of large numbers and the fact that $\max(|X_{1,n}|, |X_{n,n}|)/n \rightarrow 0$ with probability one if $E_F|X| < \infty$.

The first term in (2.5) converges to zero with probability one by Theorem 1. The second term, denoted by R_n , can be bounded as follows,

$$|R_n| = \left| \frac{n-1}{2\sqrt{n}} \sum_{i=1}^n g''(\xi_{in}) (S_{n-1}^{(i)} - S_n)^2 \right| \leq L \frac{n-1}{2\sqrt{n}} \sum_{i=1}^n (S_{n-1}^{(i)} - S_n)^2 ,$$

where $|g''(a)| \leq L$ for a in some neighborhood of S_0 , for n sufficiently large.

$$\begin{aligned}
|R_n| &\leq \frac{L(n-1)}{2\sqrt{n}} \sum_{i=1}^n \left\{ \sum_{j=1}^n \frac{M}{n^2} |x_{j,n}| \right. \\
&\quad \left. + \frac{|J\left(\frac{1}{n+1}\right)| |x_{i,n}|^2}{n} \right\} \\
&\leq \frac{L(n-1)}{2\sqrt{n}} \left\{ \frac{M^2}{n^3} \left(\sum_{j=1}^n |x_{j,n}| \right)^2 \right. \\
&\quad \left. + \frac{1}{n^2} \sum_{i=1}^n \left(J\left(\frac{1}{n+1}\right) x_{i,n} \right)^2 \right. \\
&\quad \left. + \frac{2M \sup_{0 \leq u \leq 1} |J(u)|}{n^3} \left(\sum_{j=1}^n |x_{j,n}| \right)^2 \right\} \rightarrow 0 \text{ with probability one}
\end{aligned}$$

by the moment condition on F . Hence, the theorem is established.

Naturally, the moment condition may be omitted if J trims. The analogous result on a consistent variance estimator follows.

Theorem 5: Let X_1, \dots, X_n be a random sample of size n from a distribution F obeying the conditions of Theorem 2. Let J satisfy the conditions of Theorem 2 and g have a bounded second

derivative in a neighborhood of S_0 . If

$$s_{p(g)}^2 = (n-1) \sum_{i=1}^n (g(S_{n-1}^{(i)}) - g(S_n))^2, \quad (2.6)$$

then $s_{p(g)}^2 \xrightarrow{P} (g'(S_0))^2 \sigma^2$, with σ^2 given by (2.1).

Proof: The proof proceeds by second-order Taylor expansion of the $g(S_{n-1}^{(i)})$ about S_n , and then follows the method of Theorem 2.

It should be noted that Theorem 2 (1) is (is not) a special case of Theorem 5 (4) with $g(\cdot)$ the identity, due to the identical (additional) conditions imposed on f and J in the latter theorem.

If the conditions for both Theorems 2 and 4 are satisfied, asymptotically pivotal quantities for the construction of confidence intervals for $g(S_0)$ include

$$Z_{1n} = \frac{\sqrt{n}(g(S_n) - g(S_0))}{s_{p(g)}}$$

and

$$Z_{2n} = \frac{\sqrt{n}(g(S_n) - g(S_0))}{g'(S_n)s_p}$$

If conditions for Theorems 1 and 3 hold, $g(S_n)$ could be replaced by $\tilde{g}(S_n)$ and $g'(S_n)$ by $g'(\tilde{S}_n)$ (or even $\tilde{g}'(S_n)$ if g has a bounded third derivative in a neighborhood of S_0). The question naturally arises as to which choices would be best in moderate-size sample applications of the above results. The issue of what quantity to jackknife, that is, whether to use

$$\frac{\sqrt{n}(S_n - S_0)}{s_p} \quad \text{or} \quad \frac{\sqrt{n}(g(S_n) - g(S_0))}{s_{p(g)}},$$

has not been addressed in a systematic fashion in the literature. However, the common suggestions of Miller (1974), p.12) and Efron (1972, p.1920) that the function to be jackknifed should be variance stabilized are reasonable, for example $\tanh^{-1}r$ does jackknife more satisfactorily than r , the ordinary sample correlation coefficient. Using this advice, when estimating a function g of a location parameter S_0 , the location parameter estimate S_n should itself be jackknifed to produce a variance estimator. In other words, Z_{2n} is recommended.

3. TRIAL BY NUMBERS

While the above results give a large-sample justification for use of the jackknife method for the creation of approximate confidence intervals, they leave unaddressed questions

regarding appropriateness of the technique in small-sample situations. Accordingly, a modest Monte Carlo study is in order both i) to relate the large-sample theory to samples of a size likely to be encountered in practice, and ii) to explore the behavior of the jackknife for L-statistics whose score functions violate one or more of the regularity conditions of Theorems 1-4. All computations were performed on the AMDAHL 470 V/6 at Texas A&M University.

Only location parameter estimation is considered, that is $J(u) \geq 0$ for $0 \leq u \leq 1$ and $\int_0^1 J(u)du = 1$. The distributions considered are i) $N(5,1)$, a normal with mean 5 and variance one, ii) a logistic with mean 5 and scale parameter 1, with

$$F(x) = [1 + e^{-(x-5)}]^{-1}, \quad -\infty < x < \infty,$$

and iii) $u(4,6)$, a uniform distribution on the interval (4,6). Five hundred random samples of sizes 5, 10, 20 and 40 from each of the above three distributions were examined. The score functions considered are

$$\begin{aligned}
 J_1(u) &= 0 & 0 \leq u < .05 \\
 &= 23.53 (u - .05) & .05 \leq u < .10 \\
 &= 1.1765 & .10 \leq u < .90 \\
 &= 23.53 (.95 - u) & .90 \leq u < .95 \\
 &= 0 & .95 \leq u \leq 1,
 \end{aligned}$$

a "smoothly" trimmed mean designed to obey the trimming conditions of the theorems but to fail to be differentiable at some points in $[0,1]$;

$$\begin{aligned}
 J_2(u) &= 4u & 0 \leq u < .5 \\
 &= 4(1 - u) & .5 \leq u \leq 1.0,
 \end{aligned}$$

a "triangular" weight function neither trimming nor being everywhere differentiable; and

$$J_3(u) = 6u(1 - u) \quad 0 \leq u \leq 1, .$$

a score function meeting all differentiability requirements but failing to trim.

Unfortunately, even with the above three symmetric score functions which integrate to one and symmetric parents, biases can result due to the fact that

$$\frac{1}{n} \sum_{i=1}^n J_K\left(\frac{i}{n+1}\right) \neq 1. \quad \text{Table 1 gives values of } \frac{1}{n} \sum_{i=1}^n J_K\left(\frac{i}{n+1}\right)$$

for $n = 19, 20, 39$, and 40 and $K = 1, 2, 3$. Note that a value of 1 corresponds to no bias.

(TABLE 1 ABOUT HERE)

In fact, often such biases will die out at the rate of $O(1/n^2)$ (obtained by viewing $\frac{1}{n} \sum_{i=1}^n J\left(\frac{i}{n+1}\right)$ as an application of the trapezoidal rule in approximating the integral $\int_0^1 J(u)du$ - if J'' is bounded and continuous the error is $O(1/n^2)$). Hence, the ordinary jackknife would be of little or no use in dealing with these biases. Also, a practitioner using L-statistics would doubtlessly use the modified L-statistic

$$S_n^* = \frac{\sum_{i=1}^n J\left(\frac{i}{n+1}\right) X_{i,n}}{\sum_{i=1}^n J\left(\frac{i}{n+1}\right)}$$

to guarantee unbiasedness. All of Theorems 1-5 continue to hold for these modified S_n^* if they hold for S_n , so long as $\lim_{n \rightarrow \infty} n^{3/2} \left(1 - \frac{1}{n} \sum_{i=1}^n J\left(\frac{i}{n+1}\right)\right) = 0$. Based upon the 500 samples for each sample size and distribution combination the following are estimated:

1. Bias Factors for L-Statistics

n	J_1	J_2	J_3
19	1.053	1.053	1.050
20	1.048	1.048	1.048
39	1.026	1.026	1.026
40	1.024	1.024	1.024

- a) Variance of S_n^*
- b) variance of \tilde{S}_n^* , the jackknife of S_n^*
- c) Mean of s_p^2/n , the jackknife variance estimator
- d) Mean of s_p/\sqrt{n} , the jackknife standard deviation estimator
- e-f) Percent coverages of approximate $100(1 - \alpha)\%$ confidence intervals for $\int_0^1 Q(u)J(u)du$, obtained as $S_n^* \pm t_{1-\frac{\alpha}{2}, n-1} s_p/\sqrt{n}$, where $t_{1-\frac{\alpha}{2}, n-1}$ is the $100(1 - \frac{\alpha}{2})$ percent point of a t-distribution with $n - 1$ degrees of freedom, for e) $\alpha = .10$, and f) $\alpha = .05$.

and g-h) Percent coverages for confidence intervals identical to these above, but centered on \tilde{S}_n^* , for g) $\alpha = .10$, and h) $\alpha = .05$.

Tables 2-5 present the results. It may be seen, even for samples of size 5, that the approximate confidence intervals maintain actual confidence coefficients close to the nominal 90 and 95 percent levels. Typically, the worst cases of undercoverage seem to occur for, curiously enough, the uniform parent. Intervals centered on S_n^* seem slightly better in this regard than those centered on \tilde{S}_n^* , although the

2. Results for $n = 5$

	Normal Parent			Logistic Parent			Uniform Parent		
	Score Function			Score Function			Score Function		
	J_1	J_2	J_3	J_1	J_2	J_3	J_1	J_2	J_3
a) Variance of S_n^*	.2033	.2089	.2050	.6212	.5944	.6006	.0159	.0194	.0179
b) Variance of \tilde{S}_n^*	.2033	.2375	.2132	.6212	.6250	.5966	.0159	.0258	.0208
c) Mean of s_p^2/n	.2069	.2170	.2108	.6330	.6452	.6337	.0167	.0201	.0185
d) Mean of s_p/\sqrt{n}	.4271	.4319	.4291	.7347	.7355	.7337	.1241	.1329	.1293
e) Coverage for 90% CI centered on S_n^*	.9020	.8840	.9000	.9000	.8920	.8940	.8820	.8560**	.8720
f) Coverage for 95% CI centered on S_n^*	.9440	.9400	.9460	.9560	.9500	.9460	.9280**	.9140**	.9220**
g) Coverage for 90% CI centered on S_n^*	.9020	.8840	.8940	.9000	.8860	.8900	.8820	.8200**	.8580**
h) Coverage for 95% CI centered on S_n^*	.9440	.9320	.9380	.9560	.9420	.9440	.9280**	.8960**	.9100**

** Denotes estimated coverage is at least two standard errors different from the nominal.

3. Results for Sample Size $n = 10$

	Normal Parent			Logistic Parent			Uniform Parent		
	Score Function			Score Function			Score Function		
	J_1	J_2	J_3	J_1	J_2	J_3	J_1	J_2	J_3
a) Variance of S_n^*	.0927	.1001	.0970	.3372	.3178	.3204	.0091	.0117	.0108
b) Variance of \bar{S}_n^*	.1107	.1058	.1029	.3343	.3198	.3185	.0129	.0130	.0122
c) Mean of s_p^2/n	.1032	.1094	.1054	.3315	.3203	.3176	.0083	.0115	.0104
d) Mean of s_p/n	.3119	.3176	.3138	.5511	.5401	.5402	.0899	.1037	.0992
e) Coverage for 90% CI centered on S_n^*	.9240	.8940	.9080	.8840	.8780	.8860	.8760	.8560**	.8580**
f) Coverage for 95% CI centered on S_n^*	.9580	.9480	.9540	.9340	.9340	.9300**	.9240**	.9100**	.9180**
g) Coverage for 90% CI centered on \bar{S}_n^*	.9020	.8880	.9000	.8820	.8740	.8800	.8260**	.8500**	.8400**
h) Coverage for 95% CI centered on \bar{S}_n^*	.9460	.9420	.9460	.9360	.9300**	.9320	.8860**	.9020**	.9080**

** Denotes estimated coverage is at least two standard errors different from the nominal.

4. Results for Sample Size $n = 20$

	Normal Parent			Logistic Parent			Uniform Parent		
	Score Function			Score Function			Score Function		
	J_1	J_2	J_3	J_1	J_2	J_3	J_1	J_2	J_3
a) Variance of S_n^*	.0487	.0506	.0493	.1489	.1481	.1477	.0049	.0060	.0055
b) Variance of \bar{S}_n^*	.0500	.0520	.0507	.1481	.1494	.1480	.0054	.0065	.0060
c) Mean of s_p^2/n	.0527	.0546	.0529	.1582	.1533	.1528	.0049	.0061	.0055
d) Mean of s_p^2/n	.2258	.2290	.2261	.3902	.3834	.3838	.0698	.0767	.0736
e) Coverage for 90% CI centered on S_n^*	.9100	.9040	.9120	.9000	.9040	.9080	.9000	.8940	.8960
f) Coverage for 95% CI centered on S_n^*	.9560	.9520	.9560	.9580	.9460	.9480	.9500	.9380	.9420
g) Coverage for 90% CI centered on \bar{S}_n^*	.9060	.8980	.9060	.9040	.9000	.9120	.8800	.8840	.8860
h) Coverage for 95% CI centered on \bar{S}_n^*	.9540	.9540	.9540	.9540	.9480	.9480	.9400	.9300**	.9320

** Denotes estimated coverage is at least two standard errors different from the nominal.

5. Results for Sample Size $n = 40$

	Normal Parent			Logistic Parent			Uniform Parent		
	Score Function			Score Function			Score Function		
	J_1	J_2	J_3	J_1	J_2	J_3	J_1	J_2	J_3
a) Variance of S_n^*	.0274	.0283	.0276	.0718	.0711	.0711	.0024	.0030	.0027
b) Variance of \tilde{S}_n^*	.0277	.0287	.0280	.0713	.0712	.0708	.0025	.0031	.0028
c) Mean of s_p^2/n	.0262	.0276	.0268	.0784	.0774	.0771	.0026	.0032	.0029
d) Mean of s_p/n	.1605	.1643	.1620	.2775	.2752	.2751	.0509	.0564	.0538
e) Coverage for 90% CI centered on S_n^*	.8880	.8820	.8820	.9060	.9040	.9060	.9040	.8980	.9020
f) Coverage for 95% CI centered on S_n^*	.9420	.9460	.9460	.9600	.9540	.9560	.9560	.9480	.9500
g) Coverage for 90% CI centered on \tilde{S}_n^*	.8840	.8820	.8820	.9120	.9040	.9100	.9000	.8900	.8980
h) Coverage for 95% CI centered on \tilde{S}_n^*	.9420	.9460	.9440	.9540	.9520	.9600	.9420	.9440	.9420

** Denotes estimated coverage is at least two standard errors different from the nominal.

observed differences are of the same order as the standard errors of the empirical confidence coefficients. This may be related to the typically slightly larger standard error of \tilde{S}_n^* . Consistent with the result of Efron and Stein (1979), the jackknife variance estimator appears to be typically positively biased as an estimator of the variance of S_n^* . For 28 of the 36 combinations of sample size, score, and parent population, the estimated mean of the jackknife variance estimator was greater than or equal to the estimated variance of S_n^* . Out of the 36 combinations, \tilde{S}_n^* had a larger estimated variance than S_n^* 27 times, with 3 ties (to 4 decimal places). However, the increase was typically small relative to the size of the estimates themselves. Interestingly, the estimated bias of s_p/\sqrt{n} as an estimate of the standard error of S_n^* is small, perhaps indicating that while the variance estimate may suffer from a positive bias, the standard error estimate is relatively better off. (Parenthetically, if s_p/\sqrt{n} were exactly unbiased for $(\text{Var}(S_n^*))^{1/2}$, then s_p^2/n would have a bias of order $1/n^2$, assuming $\text{Var}(S_n^*) = O(1/n)$.)

4. APPLICATIONS

The large-sample results of Section 2, bolstered by the favorable Monte Carlo results in Section 3, provide a metho-

dology for robust inference in linear models, in particular for completely randomized designs with multiple observations per treatment. We consider the model

$$Y_{ij} = \mu + \alpha_i + \epsilon_{ij} ,$$

$$i = 1, 2, \dots, t$$

$$j = 1, 2, \dots, n_i ,$$

where α_i represents the "effect" of the i th treatment. Note that we do not rule out a factorial structure for the t treatments. If, instead of the usual assumptions that the ϵ_{ij} are normally and independently distributed with mean 0 and common variance, we merely assume that the ϵ_{ij} are independently and identically distributed, symmetrically about 0, we can pursue an L-statistic approach to analysis of variance (the symmetry assumption is merely convenient - not necessary). If we assume $\min(n_1, n_2, \dots, n_t) \rightarrow \infty$ and $\min(n_1, n_2, \dots, n_t) / \sum_{i=1}^t n_i \rightarrow a > 0$, then for a symmetric score function $J(\cdot)$ meeting the conditions of the appropriate theorems in Section 2, we denote

$$S^i = \frac{1}{n_i} \sum_{j=1}^{n_i} J\left(\frac{j}{n_i + 1}\right) X_{j, n_i}$$

and s_i^2 the corresponding variance estimate, where X_{j,n_i} is the j th order statistic among those receiving treatment i . We then pool our variance estimates by

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^t (n_i - 1) s_i^2}{\sum_{i=1}^t (n_i - 1)},$$

mimicking ordinary analysis of variance. Then, an approximate

test for a contrast $\sum_{i=1}^n C_i \alpha_i$ ($\sum_{i=1}^n C_i = 0$) is provided by rejecting

the null hypothesis ($\sum_{i=1}^n C_i \alpha_i = 0$) if and only if

$$\left(\sum_{i=1}^n C_i s_i^2 \right)^2 \geq Z_{1-\alpha/2}^2 \hat{\sigma}^2 \sum_{i=1}^n \frac{C_i^2}{n_i},$$

with $Z_{1-\alpha/2}$ the $1 - \frac{\alpha}{2}$ quantile for the standard normal. An obvious modification in the critical point would permit Scheffé-type procedures. Similarly, "robust" multiple comparisons could be done. Robustness of these procedures would, of course, depend upon the robustness and convergence rates of the associated L-statistic for location - a well-studied topic. Note that the "sums of squares" for this type of analysis

can be simply computed by separately computing S^1 and s_1^2 for each treatment, and then inputting these into any standard analysis of variance package which will accept treatment means and variances as input.

5. SUMMARY

The ordinary jackknife is a computationally simple means for the construction of large-sample confidence intervals for functionals of the form $S_0 = \int_0^1 Q(u)J(u)du$. Simulation results indicate that the technique is effective for small samples.

REFERENCES

- Efron, Bradley (1979), "Bootstrap Methods: Another Look at the Jackknife," *Annals of Statistics*, 7, 1-26.
- _____, and Stein, Charles (1979), "The Jackknife Estimate of Variance," Stanford University Technical Report #120.
- Helmers, R. (1977), "The Order of the Normal Approximation for Linear Combinations of Order Statistics with Smooth Weight Functions," *Annals of Probability*, 5, 940-953.
- Loeve, Michel (1977), *Probability Theory, I*, Springer-Verlag, New York.
- Miller, Rupert (1974), "The Jackknife - A Review," *Biometrika*, 61, 1-15.
- Moore, D. S. (1968), "An Elementary Proof of Asymptotic Normality of Linear Functions of Order Statistics," *Annals of Mathematical Statistics*, 39, 263-265.
- Reeds, James (1978), "Jackknifing Maximum Likelihood Estimates," *Annals of Statistics*, 6, 727-739.
- Sen, Pranab (1978), "An Invariance Principle for Linear Combinations of Order Statistics," *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 42, 327-340.

- _____(1979), "Asymptotic Normality of a Variance Estimator of a Linear Combination of a Function of Order Statistics," *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 50, 205-221.
- Thorburn, Daniel (1976), "Some Asymptotic Properties of Jackknife Statistics," *Biometrika*, 63, 305-313.
- Wellner, John (1977a) "A Glivenko-Cantelli Theorem and Strong Laws of Large Numbers for Functions of Order Statistics," *Annals of Statistics*, 5, 473-480.
- _____(1977b), "A Law of the Iterated Logarithm for Functions of Order Statistics," *Annals of Statistics*, 5, 481-494.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER 138	2. GOVT ACCESSION NO. AD A085207	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) L-STATISTICS WITH SMOOTH WEIGHT FUNCTIONS JACKKNIFE WELL		5. TYPE OF REPORT & PERIOD COVERED Technical Report
6. AUTHOR(s) William C. / Parr William R. / Schucany		7. PERFORMING ORG. REPORT NUMBER 138
8. PERFORMING ORGANIZATION NAME AND ADDRESS Southern Methodist University Dallas, Texas 75275		9. CONTRACT OR GRANT NUMBER(s) N00014-75-C-0439
10. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research Arlington, VA 22217		11. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS NR 042 280
12. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) 12 43		13. REPORT DATE 11 23 May 1976
		14. NUMBER OF PAGES 40
		15. SECURITY CLASS. (of this report)
		16. DECLASSIFICATION/DOWNGRADING SCHEDULE
17. DISTRIBUTION STATEMENT (of this Report) This document has been approved for public release and sale; its distribution is unlimited. Reproduction in whole or in part is permitted for any purpose of the United States Government.		
18. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report) 14 TR-138		
19. SUPPLEMENTARY NOTES		
20. KEY WORDS (Continue on reverse side if necessary and identify by block number) Confidence intervals; Jackknife; Linear combinations of order statistics; Variance estimation.		
21. ABSTRACT (Continue on reverse side if necessary and identify by block number) The behavior of linear combinations of order statistics (L-statistics) under jackknifing is discussed. The asymptotic behavior of a jackknifed L-statistic is produced, and the pseudo-value based variance estimate is shown to be consistent under moderate smoothness and a trimming condition of the weight function. The results extend easily to functions of L-statistics and thus answer a question posed in Miller (1974). Monte Carlo results support small-sample applicability of the large-sample results for the construction of approximate confidence intervals.		

DD FORM 1 JAN 73 1473

EDITION OF 1 NOV 65 IS OBSOLETE

S/N 0102-LF-014-6601

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

404198

JLM